UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR

NATURAL USER INTERFACES FOR DISTRIBUTED REALITY

Sergio Serra Sánchez
Director: Redouane Kachach
Supervisor: Jesús Bescós Cano

-MASTER THESIS-

Electronics Technology and Communications Department
Escuela Politecnica Superior
Universidad Autonoma de Madrid
June 2019

NATURAL USER INTERFACES FOR DISTRIBUTED REALITY

Sergio Serra Sánchez

Director: Redouane Kachach
Supervisor: Jesús Bescós Cano

Video Processing and Understanding Lab
Informatics Engineering Department
Escuela Politecnica Superior
Universidad Autonoma de Madrid

June 2019

Resumen

El objetivo principal de este trabajo de fin de máster es estudiar e implementar interfaces de usuario para entornos de realidad distribuida. Todo ello afrontando el reto que se presenta en estas nuevas aplicaciones virtuales: no usar ningún dispositivo que ocupe nuestras manos. Esta nueva experiencia de realidad ha sido diseñada por Nokia BellLabs Spain. En este entorno DR, el usuario es introducido en un entorno virtual o remoto capturado y transmitido en tiempo real, siendo capaz de ver sus propias manos e interactuar con interfaces físicas de su entorno local. Esta tesis tratará de resolver sus limitaciones y extender sus casos de uso introduciendo la interacción con interfaces de usuario virtuales.

El punto de partida nace del estudio del "setup" ya credao por Nokia compuesto por las gafas virtuales HTC Vive y la cámara estereoscópica Zed Mini. Este sistema te permite interactuar con objetos físicos seguidos en tiempo real con códigos QR mientras observas en todo momento tus propias manos reales dentro del entorno virtual.

El objetivo clave de este proyecto será el análisis del dispositivo Leap Motion. A través de su uso se nos permitirá extender la experiencia de Realidad Distribuida introduciendo la interacción con objetos virtuales. A través de la plataforma de desarrollo Unity serán implementadas diferentes interfaces de usuario virtuales que resulten en una experiencia natural para el usuario.

Combinando ambas sensaciones de interacción se ha llevado un estudio para cuantificar el estado de inmersión de un usuario dentro de esta nueva Realidad Distribuida.

Palabras Clave

Realidad virtual, realidad aumentada, realidad mixta, virtualidad aumentada, realidad distribuida, interfaces de usuario virtuales, HTC Vive, Zed Mini Camera, Leap Motion

Abstract

The goal of this Master's Thesis is to implement a Distributed Reality (DR) environment focusing on the sensation and ability to interact with user interfaces, both virtual and real, without any controller or device. This novel experience of reality has already been designed by Nokia Bell-Labs Spain. In this approach, the user is introduced in a virtual or remote real environment being able to observe their own hands and interact with local physical user interfaces. This thesis will try to resolve their limitations and extend their user cases introducing the interaction with virtual user interfaces.

The starting point is studying to study the setup which has already been investigated by Nokia composed by the HTC Vive head mount display (HMD) and the stereoscopic Zed Mini Camera. This system allows the users to interact with tracked physical objects observing their own hands in a virtual environment.

A key goal of this project is the analysis of the Leap Motion tool that will allow the user to extend the Distributed Reality experience introducing the interaction with virtual objects. Through the Unity cross-platform game engine will be implemented some different virtual user interfaces which result in a natural experience for the user.

Mixing both virtual and real interactions we will try to quantify the grade of immersive sensation for the user in the new Distributed Reality.

Keywords

Virtual reality, augmented reality, mixed reality, augmented reality, distributed reality, virtual user interfaces, HTC Vive HMD, Zed Mini Camera, Leap Motion

Acknowledgements

Antes de nada, agradecer a Redo, mi tutor de este TFM, todo su esfuerzo y confianza depositada en mí. Sin él, las palabras que vienen a continuación no estarían escritas, al menos, sobre este papel. Hoy, entregando este TFM acaba prácticamente mi vida estudiantil. Una vida como estudiante que empezó en un colegio con la ilusión de un niño por aprender, que creció con valentía en un instituto y que se ha fraguado en la etapa universitaria. Y entre medias, miles de historias de vida, de vida real. Historias alegres, tristes, divertidas, aburridas, felices, duras... Sin duda, experiencias de vida que han definido totalmente quien soy. Y entre todas esas historias, decenas de personas se han cruzado en mi camino. Unas están, otras ya no, pero, desde luego, son los grandes artífices de hacerme estar donde estoy. Mamá, papá, Elena, no tengo palabras para vosotros. A mis padres, ya sabéis que siempre seréis mi mayor ejemplo y, sinceramente, creo que es el mayor reconocimiento que pueden recibir unos padres de un hijo, estad orgullosos. A mi hermanita, tan diferentes en tantas cosas pero tan parecidos en las más importantes, se que siempre nos tendremos el uno al otro. Familia, sois todos realmente especiales. Tía, eres una gran persona y eso, vale por encima de todo. Abuelita, tu alegría y tu ganas de vivir siempre me acompañarán. A mis amigos de toda la vida, el famoso TimetoMiller que nos unió eligió mejor imposible, juntar tanta buena gente no era una tarea fácil. Habéis sido y sois un pilar en mi experiencia de abrirme al mundo. Quiero seguir explorando la Tierra con vosotros. Mis compañeros de universidad, Pablo, Pablito, Mercedes, Clau, Ana, sois una pieza realmente incríble. Emilio, Miguel, hace dos años tomamos una decisión en nuestras vidas que creo que nos ha cambiado para siempre, para mí desde luego lo ha sido. Que un simple gesto de otro de nosotros, como la respiración, nos recuerde a estar casa creo que nos convierte en hermanos prácticamente de sangre. Mis compañeros BellCarios y compañeros de Nokia habéis sido un gran descubrimiento. Me habéis mostrado un ambiente de trabajo increíble que creía que era imposible tener. Tapia, Hugo hace tanto que nos conocemos y ahí seguimos, qué gusto da. Paula, has sido una pieza muy especial en mi vida, siempre lo recordaré como algo bonito. Y, por último, y no

menos importante, compañeros de la Cruz Roja, habéis sido un descubrimiento clave estos meses para mí. Clau, siempre con esas ganas de reir. Cris, cada día que te veo aprendo algo nuevo. Quiero seguir aprendiendo de tí y que me sigas sorprendiendo.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this first chapter, we introduced the new Mixed Reality concept of Distributed Reality and the main key differences with respect to other Mixed Realities where our novel interaction method is implemented. Then, the main goal and sub-objectives of this Master Thesis are described. Last, the document structure is explained.

1.1 Motivation

The communication between humans is our main tool to express what we feel, what we think and it is the key in the ideas interchange for our development. Throughout our history, the basis of communication has resided on personal talks when we are in the same location. The introduction of the telematics communication supposed one of the most advanced in our history and, nowadays, we have different tools such as video-calling that allow us to maintain natural conversation watching the other interlocutor despite being thousands kilometers away. Given the importance of human communication and the challenges related to improving the user experience in this topic, this master's thesis has treated to investigate new elements that enhance remote communication by using an immersive experience.

To support directly this challenge it is essential to highlight the concepts of Virtual Reality (VR), Augmented Reality (AR), and Augmented Virtuality (AV) [1] Mixed Reality (MR) which have become popular in the last years. These technologies provide a wide range of applications and nowadays we can find them in several fields such as entertainment where VR games are becoming common, education field enabling, for example, to study human anatomy in an interactive way, building design/architecture have started making use of AR and industrial environments with applications such

[1]note that AR and AV are both considered types of Mixed Reality

as assisted driving in real time makes use of MR technology.

This master's thesis is developed for Distributed Reality (**DR**) [2], a new Mixed Reality concept created by Nokia Bell Labs. This reality aims to introduce one user in a remote real (captured by a 360°camera) being able to observe their own hands and interact with objects or user interfaces (**UI**), both physic or virtual. The capacity to maintain a natural conversation watching in three dimensions your interlocutors, for example, in a corporation meeting between several users that are situated in different parts of the world, means a true immersive experience that improves any remote conversation known up so far.

Likewise, the competence to interact with physical and virtual user interfaces could open a great wide range of applications. On the one hand, sharing UIs with other users in a remote meeting could change the way of interacting with several peers. On the other hand, the capacity to manage real devices or industrial machines remotely through virtual interfaces could avoid hazardous situations in which the user could be in danger in case of being in the real place.

1.2 Goals

This master's thesis aims to investigate, implement and extend a Distributed Reality application by designing new user interfaces that enable the user to interact with virtual and physical objects within these environments in a natural way. To address the goal of this project, it has been pursuing the following sub-objectives:

1. Investigate the main concepts about the camera in the 3D world: projection 3D/2D, camera calibration, stereoscopic vision, distortion, etc.

2. Perform a deep analysis of the state of the art which involves all the realities technologies. In this way, we could include the major possible literature about Distributed Reality and we will be able to apply it to our project.

3. Become familiar and manage the new programming environments and devices used: HTC Vive VR glasses, ZED Mini Stereo Camera, Leap Motion and the Unity game engine environment.

4. Combine, integrate and calibrate the previous devices into a single one working correctly together.

5. Design and implement different user interfaces in a virtual environment that enable users to interact with virtual and real objects.

Following up the previous sub-objectives in a sequential way, it will be possible to evaluate and analyze the immersive sensation that we can feel within a Distributed Reality environment.

1.3 Master Thesis organization

The document is composed of the following chapters:

- Chapter 1: Introduction.
- Chapter 2: Study of the state of the art.
- Chapter 3: Design and development.
- Chapter 4: Demo implementation.
- Chapter 5: Evaluation.
- Chapter 6: Conclusions and future work.
- Bibliography

Chapter 2

State of the Art

2.1 Introduction

In this chapter it is included all the literature related to the project that we have studied. We have searched the most recent discoveries around each topic with the objective of recovering the most current information being able later to improve it in the development Section 3.

- First, we have explored in a general way the literature about the different types of realities. We have emphasized the advances in the communication field between two people and the capacity to interact using interfaces. As the main study, we have investigated deeply the basis and performance of the Distributed Reality concept.
- Then, it has been analyzed the three main devices used in this project: their features, their modes of use and their limitations, always oriented at our application in the Section 3.
- Finally, it has been analyzed the User Interfaces that have been developed inside the virtual and augmented realities. We have emphasized the study of the device less user interfaces in which you can interact with your own free hands.

2.2 Distributed Reality

To understand the proposal of this thesis it is essential to clarify the difference between the different types of realities defined so far and their advances into the communication and interaction field.

- **Virtual Reality (VR)**: it is an interactive computer-generated experience taking place within a simulated environment. All the communication efforts in this field have resided in introduce some virtual avatars in the same virtual environment being able to interact through voice [3].

- **Augmented Reality (AR)**: it is an interactive experience of a real-world environment where the objects that reside in the real-world are "augmented" by computer-generated perceptual information. As explained in [4] and [5], the AR technology has been used to enhance human-to-human interactions through augmented shared UI. Some users situated in the same location can interact with the same UIs using AR glasses.

- **Augmented Virtuality (AV)**: it is an interactive experience of an immersive environment that is augmented by objects from the real world. This is possible due to the video-see through capabilities of the latest HMD or by attaching an egocentric camera to the headset. An important discovery in this field has been the remote interaction among several doctors in a remote surgery [6].

As a novel conception of reality, Nokia BellLabs has developed the concept of Distributed Reality [2]. Unlike Augmented Virtuality, DR is merging different realities with the purpose of enhancing human communications and sharing experiences. At least two realities are requires, a remote one captured with a 360 camera, an a local reality captured with an egocentric camera. By selecting human body parts from the local reality [7], the DR allows the user to see himself in the experience. Concerning interaction within DR, there will possible to interact with both physical and virtual objects.

Figure 2.1: Distributed Reality examples. In a DR application, the user using VR HMD is introduced in a remote reality environment being able to observe their own body. Left image: the DR client is present in a remote meeting with real people filmed by a 360°camera. Right image: some users could meet in a virtual space like a dining room to eat together.

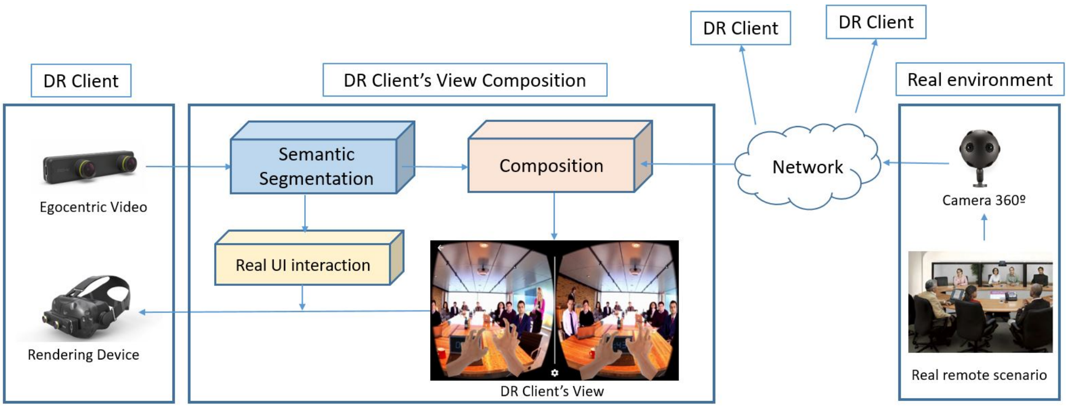

Figure 2.2: Distributed Reality Architecture. The DR architecture is composed of three main blocks: DR client, DR client's View Composition and the Remote real environment.

In this reality, the perception sought is equivalent to "being there", which means that for the user, the perception provided by her senses is indistinguishable from what they would provide with a physical presence at that space.

The main performance architecture of a Distributed Reality system is shown in the Figure 2.2:

- **DR Client**: the user setup is composed of an egocentric camera and a VR HMD which allows us to capture our own hands and display the new reality respectively.
- **Real environment**: the real space is filmed by a 360°camera capturing the perspective that could have the user if he were there.
- **DR Clients View Composition**: mixing the local semantic hands segmentation and the remote filmed environment it is created a new reality for the user in which he feels that are immersed.

In this way, the immersive experience is the key to feel a real natural experience. The 360°video or VR produce a full substitution of the user local surrounding as is exposed in [8]. A different reality is presented to the user eyes and ears in a fully immersive way: when the client moves its head, the image presented in the Head-Mounted-Display (HMD) (see cap 2.3.1) is consistent with the new pose and position of the user as shown in Figure 2.5.

The previous concept combined with the possibility of observing the parts of our local surrounding (in our case, our own body) that we want in the immersive

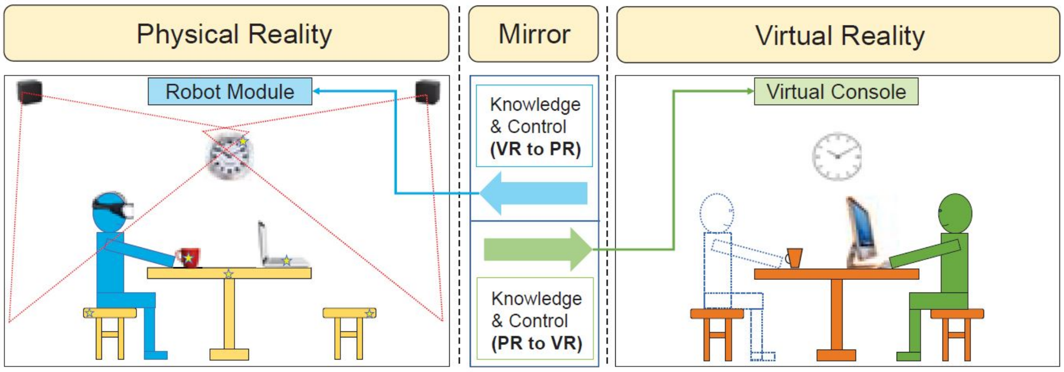

Figure 2.3: Inverse Virtual Reality example. Overview of inverse virtual reality system. The left figure is the physical reality and the right figure is the virtual reality. The knowledge and control information can be transmitted in both directions (Figure extracted from [1]).

environment, as quoted in [9], result in the concept of Distributed Reality as we have observed in the Figure 2.1. In order to select human body parts for the DR, a hand segmentation algorithm based on a skin-detection algorithm [7] is applied to the stereoscopic ZED mini camera placed in the HMD. This performance is the base of a real immersive sensation and it is obtained by using a stereoscopic camera that simulates the human vision approach (Section 2.3.2).

Distributed Reality works in parallel with the ideal Inverse Virtual Reality (IVR) environment system [1] which contains both the intelligence-driven virtual reality and the physical reality. The key to this concept is that the knowledge and control information from one environment could be transmitted to another as shown in the Figure 2.3. Any change in the virtual reality environment has to be performed in the real physical world and vice versa.

Mixing both concepts of Distributed Reality and Inverse Virtual Reality we can extend the communication applications that mainly we could approach (see Section 4). For example, using the DR technology, a machine operator could control a complex or dangerous industrial machine in a virtual environment changing the behavior of the real machine applying the IVR technology.

2.3 Devices and programming environment

Development of HMD and Mixed Reality devices are continuously evolving and improving. The improvements introduced in the input and output devices inside each of the different realities contribute even more to the immersive sensation [10].

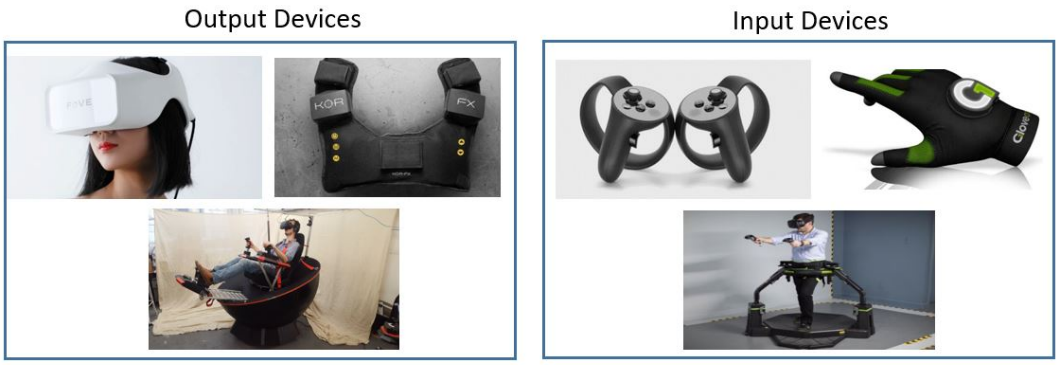

Figure 2.4: Some Input and Output devices for different realities. The left square contains three output devices which transmit to our body the stimulus to introduce us in the new reality. The more common is the HDM that is applied to our vision system (Output devices in order: FOVE Headset[1], KORF-FX[2], FeelThree[3]). The right square contains three input device which collects our movements to interact with the virtual environment. The more common is the controller which are tracking by the system simulating our hand's behavior and contain buttons to interact with virtual objects (Input devices in order: Oculus controllers[4], Gloves one[5], Owatch[6])

The more advanced device technology defined so far is collected in [1] as described below.

- **Output**: these devices provide us the sensation which we receive in our body. The more typical and extended devices are composed of stationary Head-Mounted-Display (HMD) for the sense of sight which contains accelerometers, magnetometers, and gyroscopes and use sensor fusion to combine this information with the optical tracking. However, more recent devices could transmit sensations in other ways like the trunk-body or full-body vibration as shown in the Figure2.4.

- **Input**: through these devices, the system can capture our moves or actions allowing us to interact with the immersive reality. The more typical devices are the manual controllers which are managed using our own hands. However, the creation of recent devices that detect the movement of our body (see Section 2.3.3) has opened great possibilities in the way in which we interact with the

[1]https://www.getfove.com/
[2]http://www.korfx.com/es/
[3]https://www.feelthree.com/
[4]https://www.oculus.com/
[5]https://avatarvr.es/
[6]https://www.stekiamusement.com/es/

environment. Some of these devices are shown in the Figure 2.4 like a haptic glove hand device that tracks accurately the movements of our hands.

In order to reach a high immersive DR experience, we have used jointly three main devices. Becoming each one from a different producer, it will be a real challenge merging them in only one set-up as we will describe in 3. We proceed now to explain these devices.

2.3.1 HTC Vive VR

The HTC Vive VR [1] belongs to the latest HMDs generation providing built-in head tracking, which enables an estimating position in a room-size setting [11]. They use two base stations designed to be fixed on a wall above head height, ideally more than 2 meters which contribute directly to obtain six degrees of freedom. Each base station must be placed facing the other at a maximum distance of 5 meters, and also tilted toward the HMD. It has 2160x1200 pixels for each eye visor which gives a high resolution (as shown in Figure 2.5). Besides, its wide field of view provides a certain immersive sensation.

In conclusion, its accurate position and orientation tracking [12] makes it suitable for the purpose of this Thesis.

	HTC Vive
Display	OLED
Display size	91.9 mm × 2447 ppi
Resolution	2160 × 1200 pixels
Refresh rate	90 Hz
Field of view	
Horizontal	110°
Vertical	113°
Lens type	Fresnel
Lens adjustment	IPD (60.8–74.6 mm), lens-to-eye distance ("eye-relief" adjustment)
Sensors	Accelerometer, gyroscope
Integrated camera	Yes
Audio	Microphone, jack for external headphones
Wireless	Bluetooth
Ports	HDMI 1.4, USB 3.0 × 2
Materials used	Plastic, glass, foam rubber
Weight (excluding cable)	563 g

Figure 2.5: HTC Vive Head-Mount-Display[1]. Supporting six degrees of freedom (6DOF) makes this VR HMD suitable for a Distributed Reality environment.

[1] https://www.vive.com/mx/

2.3.2 Zed Camera

In a Distributed Reality environment it is necessary an egocentric video device that captures the local reality from the end user point of view, preferably including depth information. Through this device, it will be possible to capture only our own body to introduce it into the virtual environment as explained in [13].

The best device which could contribute to this performance is the Zed Mini Camera [1]. The ZED Mini is a camera that reproduces how the human vision works. Using its two "eyes" and through triangulation, the ZED provides a three-dimensional understanding of the scene. The two cameras are separated by 12 cm estimating the depth (as applied in [14]) by comparing the displacement of pixels between the left and right images (disparity property).

As we will explain later in Section 3.4.1, the place to situate this device over the HCT Vive is determinant. The suitable position is in the center of the HMD simulating exactly the position of our eyes.

Figure 2.6: Zed Mini Camera [1]. Left Image: the Zed Camera is a stereoscopic camera which simulates the human eyes suitable for the hand's segmentation task. Center image: Zed Camera attached to the HTC Vive VR. Right image: stereoscopic vision returned by the Zed Camera.

2.3.3 Leap Motion

The proposal of this master thesis is to allow the interaction with virtual objects without using any controller. For this purpose, it is essential the correct recognition of the hand fingers and palms tracking their movements in real time. This approach was investigated first in [15] where using convolutional neural networks was obtained accurate segmented hands. From here, in [16] and [17] the performance evolved to predicts the 3D joint locations of a hand given a depth map working with high accuracy.

[1] https://www.stereolabs.com/zed/

The current main device that provides the previous technology tracking the movements of the hands with great performance is the Leap Motion tool[1].

Using two monochromatic IR cameras and three infrared LEDs, the device observes a roughly hemispherical area, to a distance of about 1 meter [18]. The LEDs generate pattern-less IR light and the cameras generate almost 200 frames per second of reflected data extracting the tracking information of the hand finger and palm. The tracking algorithms interpret the 3D data and infer the positions of occluded objects. Finally, filtering techniques are applied to ensure smooth temporal coherence of the data.

As shown in the Figure 2.7, the Leap Motion device is also attached in the center of the HTC Vive VR HMD which will be a problem to integrate along with the Zed Mini Camera in the same setup. The hand tracking works with high accuracy as we can observe in the same Figure as well.

Figure 2.7: Leap Motion Device[3]. Left image: the Leap Motion device is a small tool that use the infrared technology in the tracking hands task. Center image: the main position along with the HTC Vive is the center of the HDM. Right image: hands tracking returned by the device.

2.3.4 Unity

The workflow in this master thesis has been composed of Unity and Visual Studio. Unity is a cross-platform game engine that we have used to design our Distributed Reality environment. This platform is famous for its use in the game industry and recently it is a highlight environment in the virtual and augmented reality field.

In Unity, we have been able to design the virtual environment along with the necessary objects to manage the devices described previously. In Visual Studio we have programmed in C# the logic performance of each UI designed. Through this environment, we could have implemented the behavior of an object, for example, what happens when we push a button.

[1]https://www.leapmotion.com/

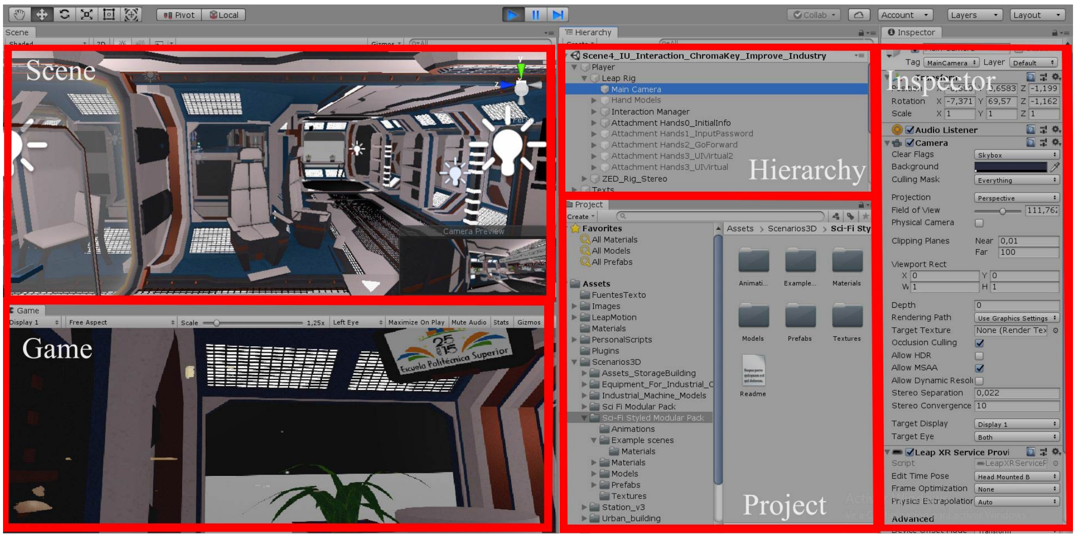

Figure 2.8: Unity engine interface. The Unity interface is composed by five main windows: scene, game,hierarchy, project and the inspector.

A simple view of the Unity interface is shown in the Figure 2.8. We can observe different tabs inside of its composition.

- The main scene tab allows us to modify manually the object's position present in the scene. We can move freely through the scene using the three axes.
- The game window is activated when we are going to play our programmed scene. In our case, every time that we want to test any change in our virtual environment, we have to wear the HTC Vive and check the good performance of the newly implemented version.
- The hierarchy shows us the objects present in our scene. As its name suggests, the objects are disposed of in a hierarchy way: inside one main object, there are more subordinated objects.
- The project window is like our library where we can find all the objects that we have imported to our project. As we will see later, each device provides its assets packets. Inside them, we can find the main necessary objects knowing as "Prefabs". These predefined objects are particularly useful due to they are composed of the essential elements to start working with one device.
- The inspector window collects the features of one object. In this tab, we can add predefined components as well as implemented scripts from Visual Studio. These components will define the behavior of one object in the scene.

2.4 Virtual user interfaces

User interfaces refer to the space where interactions between humans and machines occur. This interaction aims to allow effective operation and control of the machine from the human end, while the machine simultaneously feeds back information in return to the user [19].

As described in [20], 3D UIs are more direct or immediate. There is a short "cognitive distance" between user action and the system feedback that shows the result of that action. As shown in the Figure 2.9, a 3D User Interface is commonly composed of simple buttons, sliders, and drop-down menus. How we can interact with them and the complexity to get the action that we are trying to apply defines the utility of each User Interface.

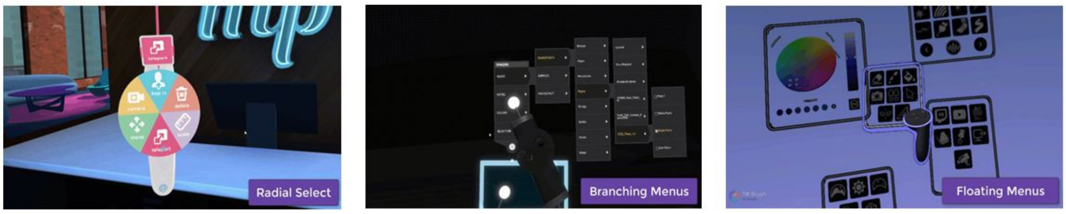

Figure 2.9: Virtual User Interfaces. The virtual UIs are composed by buttons, sliders and branch menus (Images extracted from the VR Design Review 2: Menus article[1]).

We have studied in the Section 2.3 how many input devices allow the human-machine interchange using user interfaces. In this way, the most extended devices are composed of hands controllers with buttons that simulate the hand's movements. Moving the controllers it is allowed to interact with UI as shown in the Figure 2.10. Also, there are several haptics devices like the "Manus VR Gloves" (see Figure 2.4) which are able to detect the hand pose and tracking the finger's movements.

For this reason, it is a real challenge to get the hand tracking task without using any device in our hands. In this master's thesis, we have investigated the performance of the Leap Motion tool 2.3.3 to approach this challenge: manage virtual user interfaces without using any physical device 2.4.1.

2.4.1 Contact less virtual user interfaces

We have explored the possibility of interacting with virtual objects without using any controller or haptic device. This feature could be a real progress because it allows us to keep our hands free simulating with a high self-perception the capacity of interacting

[1] https://blog.sketchbox3d.com/vr-design-review-2-menus-b0d7ddc3078

Figure 2.10: Virtual User Interfaces managed using controllers. The most common UIs are controlled using physical devices managing by our hands

with objects as we would be able to do with real physical objects.

This behavior is orientated mostly for Augmented Reality devices like the Microsoft HoloLens[1] which works with high accuracy. However, they only are able to recognize a few numbers of hand gestures which limits highly their use. As we can observe in the Figure 2.11, the HoloLens allows us to interact with complex user interfaces but only using two simple actions like press two fingers simulating click a button.

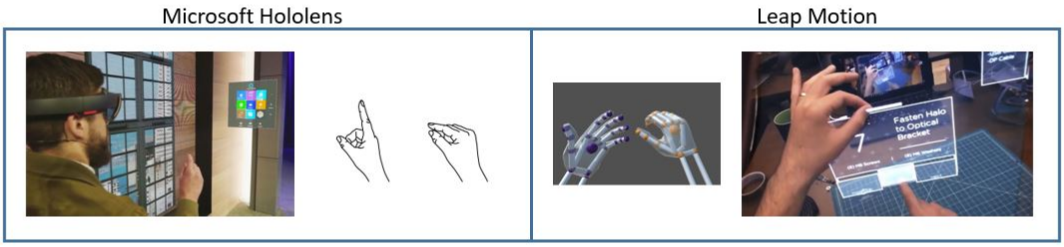

Figure 2.11: Contact-free devices interacting with UIs. Left square: HoloLens tool allows us to interact with virtual objects in an Augmented Reality environment. Right square: the Leap Motion tool allows to recognize our hand in both Virtual and Augmented Reality environments. Perform with high accuracy and contributes to greater freedom to the UIs interaction.

Following the same path, this master thesis has investigated several tool that allows us to keep our hands free and interact with virtual UIs in a Virtual Reality environment. The chosen device has been the Leap Motion tool. As we have commented in 2.3.3 and as we will study in the next Chapter 3, the Leap Motion tool provides high tracking accuracy to interact with virtual UIs. As we can observe in the Figure 2.11, the differences with the HoloLens are notorious because of the higher performance of Leap Motion allowing to interact in many ways with virtual objects. Although the investigation in this topic is not so extended and the main examples that

[1]https://www.microsoft.com/es-es/hololens

we could find are recovered in the Leap Motion Documentation[2] [21], this master's thesis has tried to recover and extend their use as far as possible.

[2]https://leapmotion.github.io/UnityModules/

Chapter 3

Design and Development

3.1 Introduction

This chapter describes how from the investigated material and the tools studied in the state of the art, it has been developed and implemented a Distributed Reality environment in which the user is able to interact with virtual user interfaces. The basis of this project is to integrate the new Leap Motion device described in 2.3.3 to recognize the user's hands beside the two tools used so far (Section 2.3.1 and 2.3.2). For this proposal, a previous exhaustive study of their performance and the calibration of the three devices will be a real challenge to get our objective. Once that, we will investigate the possibilities that provide this advanced device creating new virtual user interfaces that will be designed in 4 and tested later in 5.

3.2 General system proposal

This section describes the proposed system which introduces the Leap Motion tool in a DR environment. The proposed setup is reflected in the block diagram shown in the Figure 3.2.

First, we have started from the main Distributed Reality architecture (described in the state of the art section, Figure 2.2). In this general block we introduced the new Leap Motion device pursuing the goal of this thesis: interact with virtual user interfaces in a natural way. As we can observe in the Figure 3.1, the physical device is introduced in the DR Client setup along with the other hardware. As we will see in 3.3.3, this device provides the necessary tools to interact with virtual UIs naturally.

To organize its integration and to take full advantage of its possibilities, the new module has been studied sequentially as we can see in the Figure 3.2. Each proposed

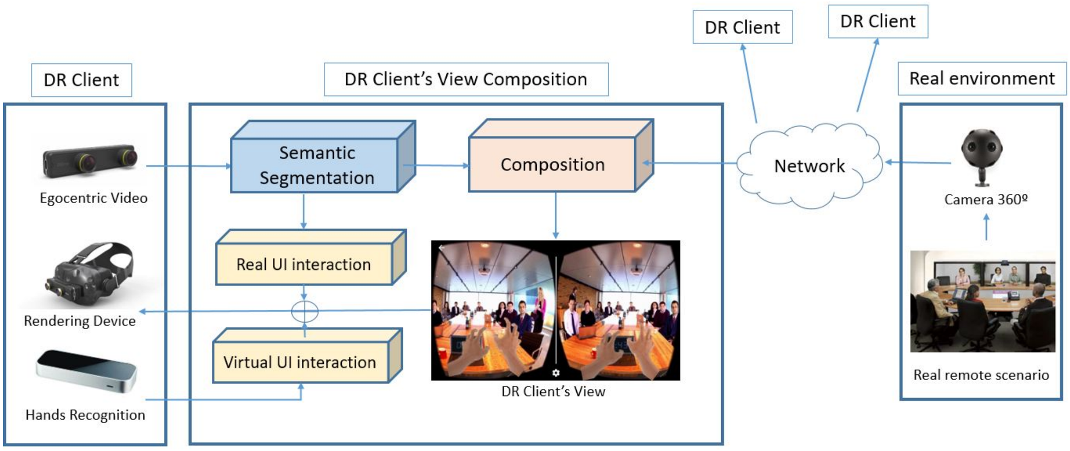

Figure 3.1: Extended Distributed Reality Architecture. The main DR architecture proposed up so far allowed us to interact with physical virtual objects. The new Leap Motion device is introduced in the DR Client setup and provides support to interact with total virtual objects.

block or module is developed, explained and detailed referencing the corresponding state of the art described in 2. For each model, it is described the issues raised to propose below their solutions.

1. **Leap Motion device**: We start from the Leap Motion device described in 2.3.3. The Leap Motion's software and hardware platform bring your bare hands directly into virtual and augmented reality. Through the "Orion beta SDK" provided by their developers "Project North Star", it is possible to check its initial performance. As we can observe in the Figure 3.3 the device tracks our hands in each frame using their infrared cameras. Also, the system allows us to interchange the model of the represented virtual hands.

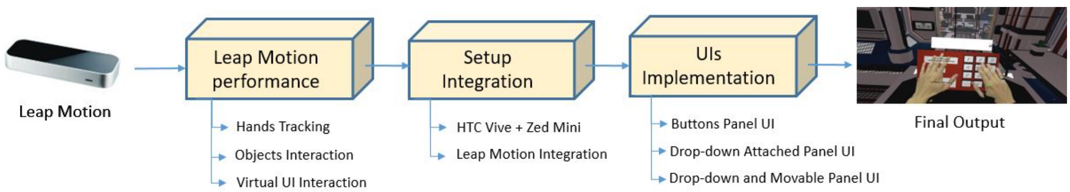

Figure 3.2: UI Virtual Interaction Workflow. Starting from the Leap Motion device we have sequentially studied its performance. First, it has been implemented simple cases of use of the Leap Motion, then, the device has been attached to the setup defined up so far. Finally, it has been implemented some different UIs.

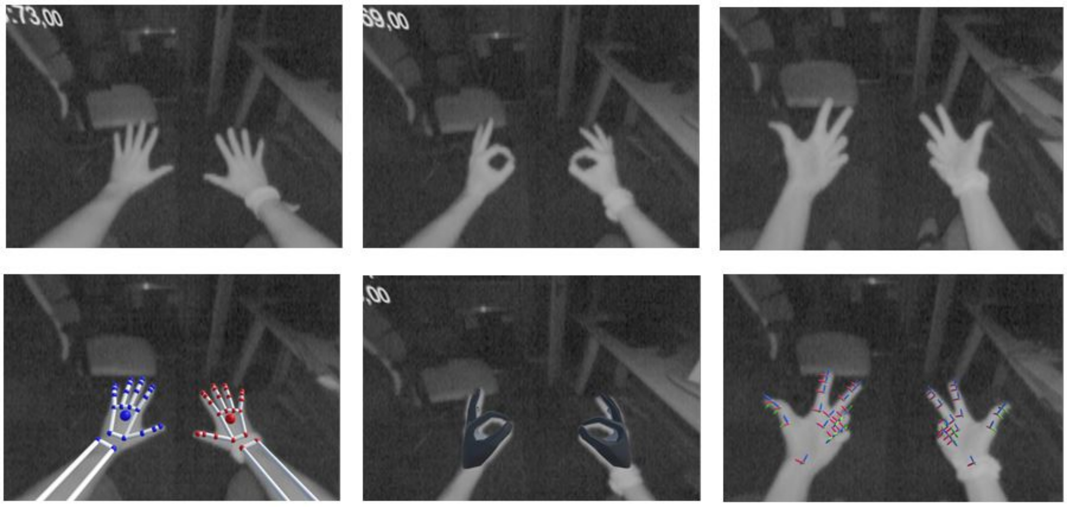

Figure 3.3: First Leap Motion experiments. The Leap Motion device using their infrared cameras allows tracking with high accuracy the hand movements. In order to observe the tracked hands in different ways, it is possible to interchange the represented hands models.

2. **Leap Motion performance**: after the first contact and some basic experiments with the device, we have extended its study to cover all the applications that we could implement using it. For this, we have implemented some simple case of virtual objects interaction such as taking, dragging and throwing objects, pushing buttons and interacting with sliders.

3. **Setup Calibration**: in this section we analyze how and where could be attached the Leap Motion device to the HTC Vive HMD (see Section 2.3.1). This integration has to be as accurate as possible to increase the immersion sensation and extend the interaction capacity. The main goal followed is that the video hand captured by the ZED Camera was exactly the same that the hand tracking by the Leap Motion device.

4. **UIs Implementation**: making use of the calibrated setup, we could design virtual User Interfaces. Starting from the UIs provided by Leap Motion' SDK, we have designed and modified them to create several new virtual UIs that allow us to interact with the virtual environment.

5. **Final user output**: finally, integrating the previous modules we obtain the final user output for a Distributed Reality client in which we are able to interact with virtual user interfaces.

3.3 Leap Motion performance

In this section is detailed the Leap Motion performance. For this, we analyzed the behavior of the device in several situations oriented to interact with virtual UIs. For this purpose, we implemented the following simple use case applications:

3.3.1 Hands tracking

Using the Unity Core Assets 2.3.4 provided by the Leap Motion developers we can test directly the hands tracking performance. Dragging the main 'Leap Rig' prefab it is possible to observe in our HTC Vives glasses (see 2.3.1) as our hands are recognized and represented virtually.

As we can observe in the Figure 3.4 the hands tracking works with high accuracy. When we move rapidly the hands, their corresponding virtual hands react in the same way. Testing the fingers movements it is possible to observe as they interact well between them in case of simple occlusions.

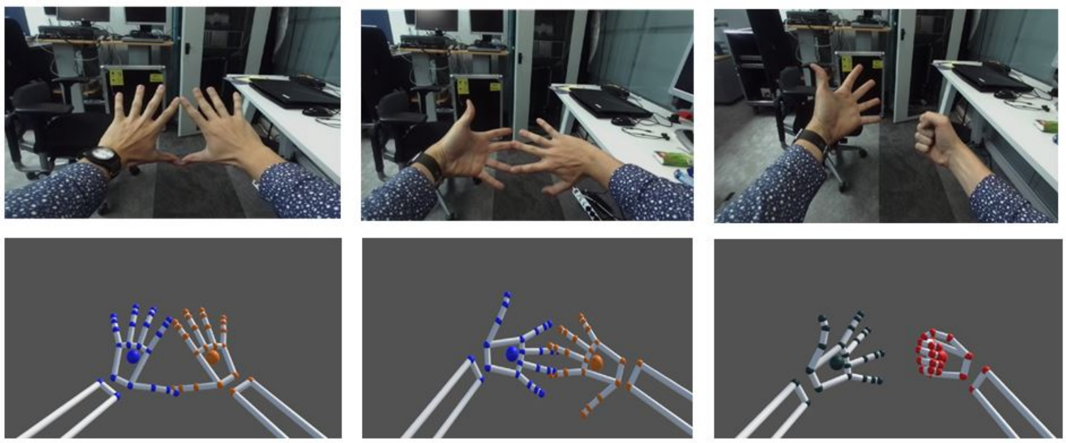

Figure 3.4: Hands Tracking Performance. First row: the real hand's gestures. Second row: the virtual recognized hand by Leap Motion. We can observe as the hands tracking perform with high accuracy even when we interact between the different hands or fingers.

One important case we studied is the recognition when we occlude any part of the hand. If we cover one hand with an object or with our other hand, the system is not able to detect the occluded hand (see figure 3.5). Also, there are important bugs that could affect our system when we occlude our hand with the same hand. As we can see in the Figure 3.5 the stretched finger that is partially occluded by our hand is bent in the virtual representation. As we will observe in 3.3.3 this can be a problem when we touch any panel located in front of us with buttons or sliders.

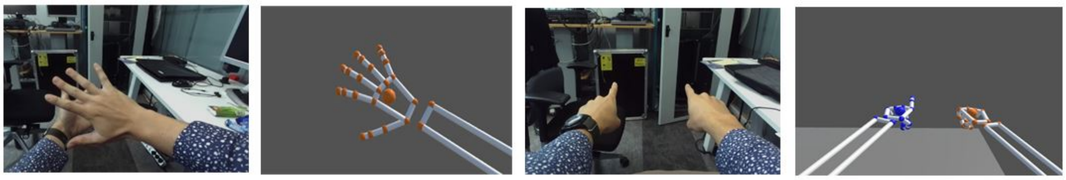

Figure 3.5: Hands Tracking Issues. There are two main cases in which hands tracking do not work well. Left image: When we occlude a hand entirely, the device is not able to recognize the hand. Right image: if the view perspective is far away, the occluded fingers are not well tracked. In this image, the left stretched finger is detected correctly, however, the right finger is bent.

3.3.2 Objects Interaction

Virtual objects interaction is the basis of using Leap Motion. This device allows us to interact with virtual objects that are within reach in our scene. Through it, we can touch objects and displace them, take and grasp objects when we close the hands and throw the objects when we open the hands again. These movements and interactions cover a bug range of the possibilities that a human could do with its environment and open a high range of applications.

To test its performance we have implemented some simple use cases as we can observe in the Figure 3.6.

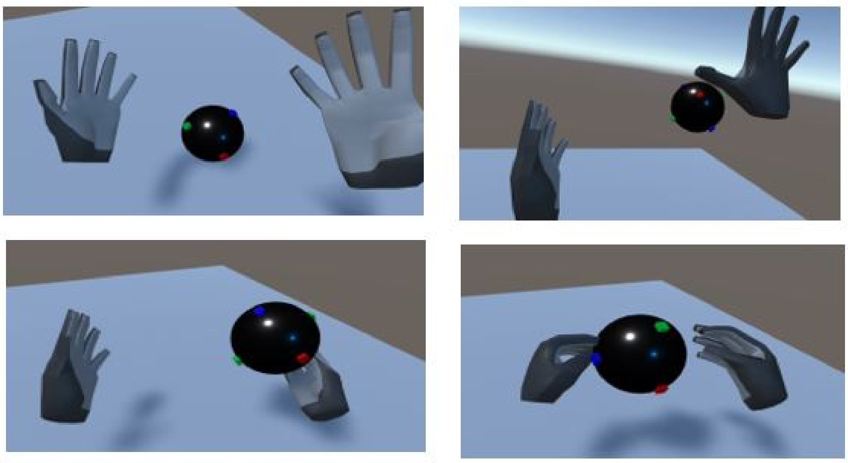

Figure 3.6: Object Interaction Performance. The Leap Motion tracks our hands allowing us to interact with virtual objects. The accurate tracking allows detecting when we have an object grasped even when we change the object of hand (bottom row images). When we open the hand, we can throw the object in the moving hand direction.

We can touch any object fixed as 'RigidBody' in the Unity environment allowing the user to move it. Using the device, as explained in [22], if we touch the object moving the hand rapidly, the object goes far away considering the velocity of the hand. If we program them also with the 'InteractionBehaviour' feature, we can take and drag them along the virtual scene when we close the hand near to the object. As we can observe in the Figure 3.6, when we open the hand again the velocity of the hand is calculated and we are able to throw the object.

3.3.3 Virtual user interfaces interaction

Interacting with virtual User Interfaces with our real hands is a real challenge to interact with a virtual environment in a natural way (as we have seen in 2.4). In this sense, Leap Motion device enables us to interact with user interfaces composed by buttons and sliders as an extension of simple objects.

In the Unity environment applying to the virtual objects the 'InteractionButton' or 'IteractionSlider' features, we can process them as physical objects. This generates an event whenever a force is applied on the virtual object. As we can observe in the Figure 3.7, Leap Motion developers provide several simple UIs that allow a great number of possibilities of interaction. From simple panels with buttons and sliders until real complex objects attached to our hands.

The buttons interaction works great in the major of cases, although, as we saw in 3.3.1, when we try touching a button with a stretched finger, sometimes the recognition is not perfect and it could be hard to interact with it.

For this reason, we **propose two solutions** to improve these sensations and make more reliable the capacity of interaction.

1. Place the UIs under of our point of view. It implies situate the UIs to a height in which we avoid to occluding our stretched finger with our own hand. As we can observe in the Figure 4.2 when we place the UIs in lower height, the finger is maintained stretched.

2. Push or tap buttons and sliders with two or more fingers. Although using the previous proposition we improve the performance, sometimes the accuracy is not enough (as we will observe in 3.4) to recognize a finger pressing a button. For this reason, as we are able to observe in the Figure 4.3 we improve the performance using two or more fingers being our sensation more accurate.

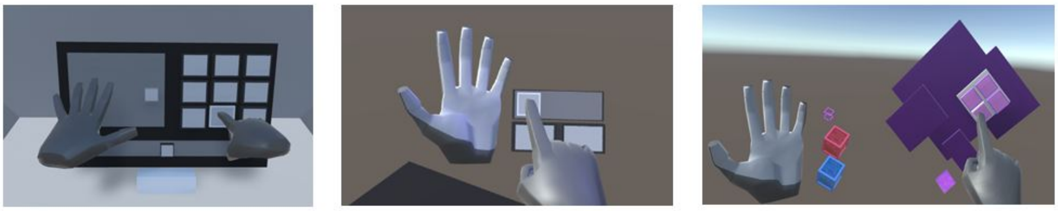

Figure 3.7: Virtual UIs Interaction. The Leap Motion software provides for the Unity environment some predesigned UIs. From simple panels composed of buttons until more complex attached boards with sliders.

These prototypes will be really useful as we will be able to see in the Section 3.5 to create and configure our own functionality. Setting each button and slider to trigger an event, we will be able to modify each UI to our needs. Through these interactions, we will be able to induce changes in the virtual environment interacting directly with it.

3.4 Setup integration

One of the challenges we faced in this thesis has been the correct integration of the three devices used for a DR view's client. The search for a good performance among the three devices working in parallel has been an essential task to obtain good results.

Their developers propose how to integrate their devices into the HTC Vives. However, introduce both devices into the same setup is not explained. As we will see below, the key is introducing an offset that fixes the displacement of change the original position that it would have.

3.4.1 HTC Vive integration with ZED Camera

The structure of the setup before introducing the Leap Motion device was composed of HTC Vive along with the ZED Camera. This setup has already been used in other applications as mentioned in [23] where it is employed in an immersive environment for medical therapy or in [24] where it is applied to ensuring safety in risk vehicle situations.

The position of the ZED Mini Camera over the VR glasses has to be exactly in the middle of the device simulating the eyes view. This integration has to be accurate to ensure that the hands we see are situated exactly in the position if we did not wear the HMD.

Figure 3.8: ZED Mini Camera setup. The Zed Mini Camera captures the local user environment in a stereoscopic way. For this reason, it can be displayed to the user in the binocular HTC Vive HMD (left image). The two cameras simulating the human eyes provides the capacity of calculating the depth of applying the disparity property.

For this purpose, the 'Stereo Labs' (ZED Mini Camera developers) provide the mount to attach it along with the HMD HTC Vive just in the center (see Figure 2.3.2). As we are able to observe in the Figure 3.8, the ZED camera film the real environment in a stereoscopic way and it is displayed in the HTC Vive to our eyes.

This setup is the base of three main features of our DR system:

1. Integrate virtual objects in our real environment. While the Zed Mini Camera captures the real environment around us, the HTC Vive along with their accurate tracking establishes the virtual objects static over that real environment. As we can observe in the Figure 3.9, introducing virtual objects in our scene's Unity, we have the sensation of these virtual objects are in our reality being in the same place when we move the HMD.

2. A great feature of the ZED Mini Camera is the capacity of observing the world in depth as we have commented in 2.3.2. Using the disparity property, we can consider the real environment captured by the ZED Camera and occlude the correspondent virtual objects. In the Figure 3.9 we can observe this important approach in which if we put my hand ahead of a virtual object (as a UI), this object is covered and it is not shown to our view in the HMD.

3. Applying a chroma key to observe only our own hands. Given that we want to get a Distributed Reality environment in which the background is composed of a real remote environment captured by a 360°camera or a virtual context, we are interested in capturing only our own hands with the Zed Mini Camera. For this reason, apply a chroma key around the skin color's tones is a simple but effective way to get it. We can observe this approach in the Figure 3.10 in which the user is introduced in several virtual environments observing their own hands.

Figure 3.9: ZED Mini Camera virtual objects integration. Left image: through the mixed HTC Vive and Zed Mini camera setup we are able to integrate virtual objects in the real world keeping them statics in the same location. Centered image: the depth property of the Zed Camera allows us to treat the virtual objects as real objects. When we cover a virtual object with our hands, we are not able to observe it. Right image: when our hand goes through a virtual object, our hand disappears.

4. Introduce virtual objects in our scene using QR patterns in physical objects. Given that the Zed Camera is composed of two independent cameras, Nokia BellLabs use the left one to detect QR patterns in our real environment using the Aruco library. This library tracks defined markers previously. When we move the physical objects with the QR pattern printed, Aruco recalculates the 3D position and place it correctly in the scene.

Figure 3.10: ZED Mini Camera Chroma Key. Applying a chroma key around our skin color, we can observe only our hand in a virtual environment.

We can observe an example of this in the Figure 3.11. Following the main marker is represented a video in the center of the square. When we tap the other QR patterns at the bottom of the panel, we are able to interact with the video. In the evaluation section, we will analyze the differences between this kind of interaction and the proportioned by Leap Motion only with virtual objects.

3.4.2 Leap Motion Integration

'Project North Star' (Leap Motion developers) proposes how to integrate its device in the HTC Vive. The ideal place is in the middle of the HMD (shown in 2.3.3) as

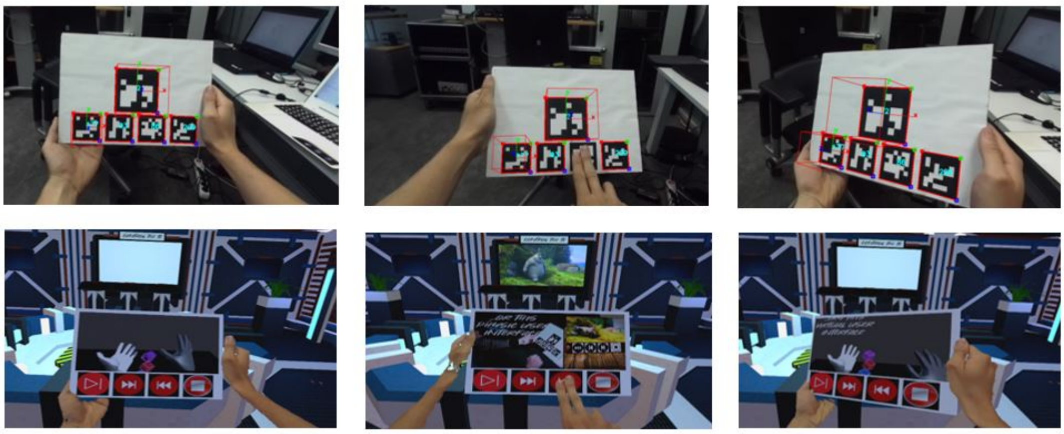

Figure 3.11: ZED Mini Camera QR Patterns. In a Distributed Reality environment defined before introducing Leap Motion, the user was able to interact with physical objects composed of QR patterns. Using the virtual tablet (which is backed by a physical light cardboard) the user can interact with the virtual scene.

it was proposed for the ZED Camera 3.4.1. In this position, it is possible to ensure that the recognized hand by the device are displayed exactly in the same position in the HMD. However, as that place is occupied by the ZED Mini Camera, it has been necessary to find another position to put it.

The **proposed solution** in this thesis (interchanging opinions with the Stereo Labs support center) has been to put the Leap Motion device just below of the Zed Camera calibrating it later in Unity. For that, it was necessary to get the mount of the device. Using a 3D printer we could get the piece rapidly printing it in only a few minutes. As we are able to observe in the Figure 3.13, we get the final setup with the three devices in the same piece.

With this setup, it was necessary to carry out the calibration process in which was applied a compensation to place the real and virtual hands in the same position. This offset parameter was introduced in the 'Zed Camera Prefab' being dependent on the 'Leap Motion Prefab'. When we change this offset parameter through the x-y-z-axis, the recognized virtual hands are displaced. By trial and error, we tried to find the best approximation that maximizes the precision among both virtual and real hands. As we can observe in the Figure 3.12, without the offset implementation, the virtual hands are uncompensated. After applying the offset through the x-y-z-axis, we obtained both hands in the same 3D place.

This final setup combines all the features that we have looked for a Distributed Reality environment which improves the user immersive experience.

A user that uses this setup will be able to see his own hands in a remote real or

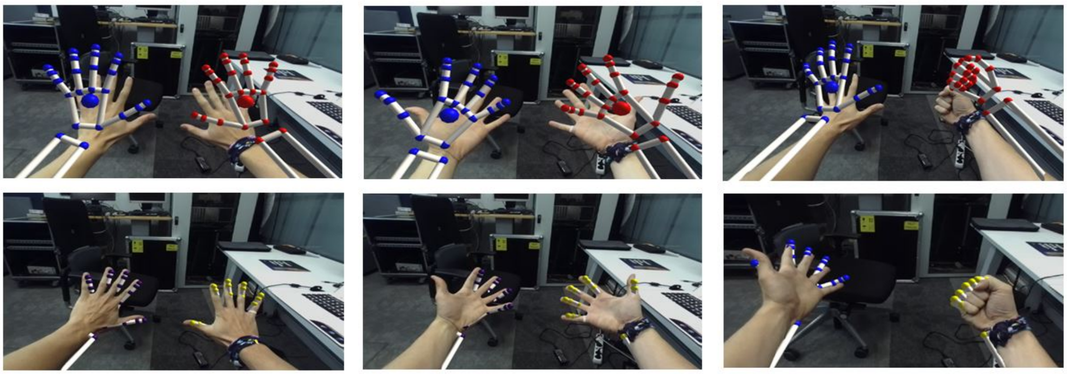

Figure 3.12: Leap Motion calibration. First row: the virtual hands are uncompensated not matching with the real hands. Second row: after applying an offset, the virtual hands match with the real hands. Now, we can interact correctly with virtual objects.

virtual environment. He can observe and interact with physical virtual objects and entirely virtual pieces, specifically, with user interfaces that could be in the environment or attached to their hands as we will observe in 3.5. As we could observe these features open a wide range of possibilities (as we will observe in the next Section 3.5). It will change completely the way in which we interact with a virtual environment.

Figure 3.13: Final setup composed of the HTC Vive VR, Zed Mini Camera and Leap Motion in one piece. The Leap Motion device is placed under the Zed Mini camera. This final setup, applying the explained calibration, allows us to interact correctly with virtual objects.

3.5 UIs implementation

The final module to complete the Leap Motion analysis is the UIs's design and implementation block. Through this study, we have determined which UIs could be useful in a DR environment. To do that we have focused our investigation and logic programming in three different UIs.

3.5.1 Buttons panel UI

Subjectively, we could say that the simplest UI is composed of a panel that contains interactive buttons represented in the Figure 3.14. A simple use case could be the introduction of digits when we push one button, for example, introducing a password.

To test this use case we implemented this logic as Unity C# script. When we push a button of the panel, a text is modified according to the tapped button. As if that text were a password, when the string is equal to a predefined text, an event is fired in our virtual environment. Figure 3.14 shows an example of this use case. When we introduce the correct combination of numbers, the sphere in front of us disappears immediately.

As we will see in 4 in 4 this kind of UI could be really useful in a DR environment opening the possibilities of interaction with the virtual environment or with other DR users showing them any virtual information of interest.

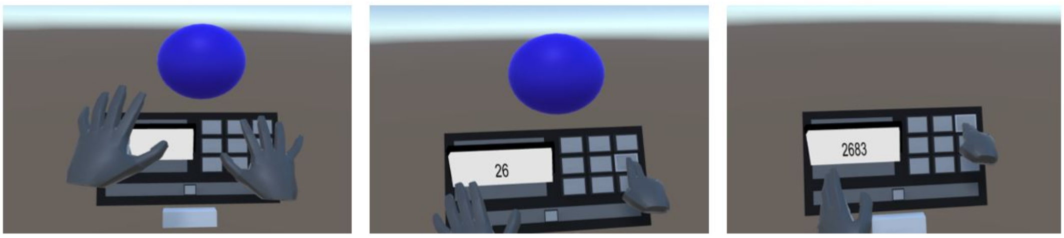

Figure 3.14: User interfaces composed of buttons. The images show the sequence of a user introducing the password. When the correct password is introduced the blue ball disappears from the scene.

3.5.2 Drop-down attached panel UI

Leap motion SDK comes with an interesting UI concept (which in fact was designed to be used in Augmented reality scenarios). This UI could be applied in DR environments as well since we are able to see our real hand. It is related to an attached UI that appears only when we rotate the left hand upwards. As we can observe in the Figure 3.15, this is an interesting concept because we can have access to this UI whenever we want only turn our hand and it can provide us at all the time the information we need. We can move along this information using the buttons panel which is shown when the UI appears as well.

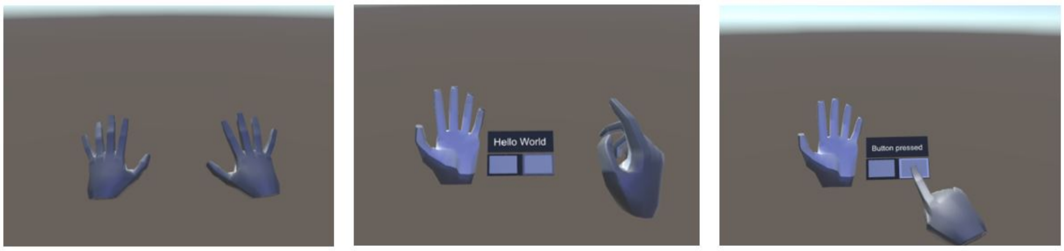

Figure 3.15: Drop-down attached panel. The three images show the process of interaction with an attached UI to our hands. When the left hand is upward, an attached UI appears. We can interact with the two buttons changing the text showed.

As a simple example of use, we have implemented the logic of this UI showing different information for a DR client. As it showed in the Figure 3.15, when the users turn their left hand, a panel appears showing a text. At the bottom of the panel, two buttons allow the user to interact with that information going forward or backward of the showed text.

We will be able to observe as well in 4 as this kind of UI could be really useful in a DR environment to keep the user informed all the time.

3.5.3 Drop-down and movable panel UI

Finally, another useful UI that has been already commented in 3.3.3 is a drop-down and movable panel. As we can see in the Figure 3.16, this UI provides us the capacity to display a drop-down UI with buttons that previously was attached to our left hand. As the previous UI, the user does not see anything until he rotates the left hand upwards. Then, the user can take a simple cube and throw it wherever he wants. From that cube, a button panel appears and the user is able to interact with it.

This performance gives the user great freedom to situate the drop-down UI wherever he wants. It could be really interesting to have more complex UIs only when it is necessary.

As a simple case of use, we have implemented one of this UI to modify one 'GameObject' in a Unity scene when a button is pushed. Figure 3.16 illustrates this concept. The user takes and drags the cube displaying the drop-down UI. As we have implemented in the first UI 3.5.1, when the user pushes one button, an object disappears in the Unity scene.

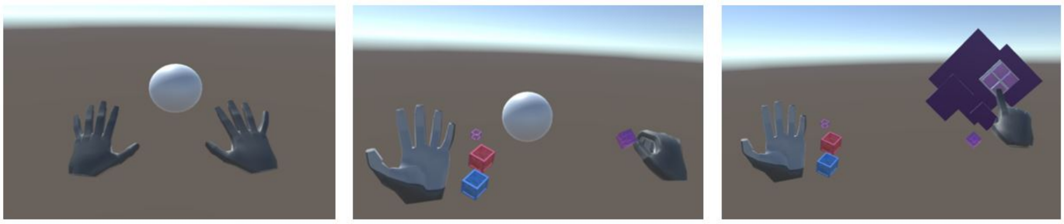

Figure 3.16: Drop-down and movable panel. The three images show the process of interaction with an attached UI to our hands. When the left hand is upward, an attached cube appears. If we take one of them and throw it, it is unfolded a drop-down menu. We can also interact with the buttons. When one button is pressed the sphere disappears.

Chapter 4

Demo implementation

4.1 Introduction

This chapter describes the develop of a "Virtual User Interfaces Demo" that has been implemented to collect all the advances implemented in the previous chapter 3. The concept of a "Demo" in the industry field is a really important element to showcase some idea or a new technology. When designing a demo several elements have to be taken into account but the most important one must be creating a new experience for the user.

The main goal of this demo is to provide the user with a new immersive experience. In addition to sense and feel the virtual environment, he can feel and see his body (in this case, his own hands) while immerse in a virtual scene. Given that the designed UIs in the previous chapter were so simple, we tried to make them more attractive and striking to the human vision. The virtual environment (detailed in 4.1.1), is also an important element in the process of design a Demo and it has been chosen carefully.

The Demo implementation has been divided into three stages in which the user has to interact with a different user interface in each one. The demo was designed to be self-explanatory. Thus, user shouldn't need any external help besides the information provided to him through the Demo.

4.1.1 Virtual environment and style design

First, we choose the virtual environment which involves our Demo. Given that creating a new virtual environment from scratch is not part of this thesis goals, we used an already designed environment. The Unity Store provides many virtual objects. From single simple virtual objects like a chair to complex models of virtual environments like an industry scenario.

Figure 4.1: Two virtual environments tested. The corner images show two of the virtual environment tested. Finally, we choose the Virtual environment 2 composed of modern furniture. The center image shows the adapted local environment for a DR client. This homogeneous color surrounding the user improve significantly the chroma key applied in the hand segmentation.

As we have shown in 2.3.4, the Unity objects already designed are grouped in Assets packets. At the time of chose one virtual environment, several of them are available for free. In this process, we downloaded and tested some of them related to the industry or modern environments (as shown in the Figure 4.1). Finally, we choose a modern home with futurist furniture (Virtual environment 2 in the Figure 4.1) that adapts perfectly to the attractive user interfaces concept.

It is important to highlight that to improve the immersive sensation with a correct segmentation of the hands, we have used a green local environment. As we can observe also in the Figure 4.1, the user is surrounded by a green canvas to make the chroma key more effective. In this way, the extracted hands are better defined and have fewer artifacts.

Also, along the next stages, we will see the style applied to the user interfaces. We have chosen a red tone style with white buttons. These colors are more attractive and easier to see for the user as we will observe below.

4.1.2 Stage 1

In the first stage of this Demo, we try to inform the user about the goal behind the experience as much as possible. To do that, all the information is orientated to explain the aim of this thesis: interact with virtual UIs watching your own hands creating Distributed Reality environment.

The most important information in this section is a loop video showing the explanation of the UI used in this stage: a drop-down attached UI as we have shown in 3.5.2. This UI, called "The User Personal Assistant", is really useful because using it, the user is at any moment informed about what he has to do. The main idea of this UI is that whenever the user does not know how to continue through the Demo,

Figure 4.2: Drop-down attached panel UI. When the user turns upward his left hand an attached UI appears. The user can interact with the text showed progressing forward and backward using two buttons.

this UI gives him the information about what is the next step to proceed.

Figure 4.2 illustrates an example of this UI. When the user turns his left hand upwards, an attached user interface popup along with it. This UI is composed of a text, two buttons and a video that makes more visual the interface. This text informs the user about its disposition any moment during the Demo. The user is able to go progress forward and backward through the text and finally, pushing the last Next button the user is advanced to the next stage.

The logic programming implemented in C# for this UI is not so complex. Whenever we detect an interaction between our hand and the button, the text changes depending on the Next or Back button pressed. Finally, when the final Next is pressed, the 'Game Object' corresponding to the user is moved at the next position of the virtual environment.

4.1.3 Stage 2

In the second stage, the user is placed in front of a panel with buttons. Behind the panel, a door and a text with some instructions are showing the purpose of this stage: open the door introducing the correct password in the buttons panel. As we can observe in the Figure 4.3, this panel is composed of twelve buttons, different text squares and a video explaining how to use the personal attached UI.

To know the correct password, as we have said, the user has to use its personal assistant. Turning upward its left hand, the previous UI appears attached to his hand as we can observe in the Figure 4.3. This time, the user can press a button to display the password and push another one to hide it.

Knowing the password, the user can introduce in the buttons panel the corresponding number pushing each button (observe the Figure 4.3). When the password is correct, the user press the Accept button and the door is opened. Finally, the user is advanced to the final third stage.

Figure 4.3: Buttons panel UI. First and third image: show the buttons panel to introduce the correct password. The UI is composed of twelve buttons, one of them to delete and another one to accept the introduced number. To know what is your password, the user disposes of an attached UI that shows him the corresponding number.

The logic programming implemented in C# is centered on how each button influence to each text. When it is detected an interaction between our hand and any button, the corresponding text changes. Finally, when the password is correct, the 'Game Object' corresponding to the user is moved at the next position of the virtual environment.

4.1.4 Stage 3

The third stage is based on the use of a virtual drop-down panel UI showed in 3.5.3 and a virtual physical UI explained in 4. The user is placed in a futurist room in which he can observe a large TV in front of him. This stage aims to interact with this TV using the two main interfaces described: one totally virtual drop-down UI and one physical virtual UI. The idea is to compare what is the sensation of interacting in the same space with both objects, one physical and another one virtual.

At the beginning, the user observes a panel in which two ways to control the TV are shown:

1. First, turning upward its left hand, the user can take one virtual cube, as we have seen in 3.16, and throw it to deploy a floating virtual UI composed by buttons.

2. Second option: use a physical cardboard tablet (see Figure 4.4). In fact, the user sees this tablet enhanced in his virtual environment. By using Nokia Bell-Labs technology we can track the tablet in real time and use it to project a virtual UI that can be used by the user. Inside the virtual scene the user is able to press the enhanced tablet buttons as if were a physical/real one.

Figure 4.4: Drop-down movable UI and physical virtual UI. First row: shows the corresponding images of a Drop-down movable UI. The user can throw an attached cube to his hand showing a drop-down panel with buttons. Second row: shows how the user can manage a physical UI tracking their corresponding QR patterns.

These two interfaces are composed of buttons with which it is possible to manage the TV. The buttons allow us to play and pause the TV, watch the next or previous video, and stop the video.

As we will comment in the next Evaluation chapter 5, this Demo is essential in the evaluation process to measure the sensation of interaction with two totally different UIs. The feedback of subjective tests run in the Chapter 5 are important to understand the advantages/disadvantages of each approach so we can determine which user interfaced provides a better user experience.

Chapter 5

Evaluation

5.1 Introduction

This chapter contains a deep subjective evaluation of the previous demo (exposed in Chapter 4). Finally we decided to run a subjective experiment to evaluate the quality of the demo user experience since rating this kind of environments is not a straightforward task. The proposed evaluation in this thesis is based on two standardized methods to rate a subjective user experience. This approach has been already used in [25] where these two methods are employed to measure the user UIs interaction (2D and 3D) satisfaction.

The two evaluation methods are based on fill out a paper composed of different subjective questions. The first one, NASA TLX [26], is a more general model used to evaluate a work task measuring the workload employed during an activity. The second one, UEQ [27], is a fast questionnaire to review the User Experience of interaction with an interactive product.

Through this evaluation, we will try to extract the useful information about the implemented UIs. This report will allow us to conclude what are the best skills of our UIs and what are the points to improve in the future.

5.2 Participants

A total of 10 participants (8 males and 2 females, age = 20-60, mean = 38) participated in this evaluation process. These members belong to the Nokia BellLabs company and have been inquired with total anonymity.

In order to know what is the level of knowledge by the users in the field of virtual/augmented/mixed reality, the participants have been asked about this topic as

	None	Novice	Frequent
Virtual Reality	0%	20%	80%
Augmented Reality	10%	20%	70%
2D Virtual UI	20%	10%	70%
3D Virtual UI	20%	20%	60%

Table 5.1: Knowledge of the users in the Distributed Reality field. Most of the users know these environments. Most of them are unfamiliar with 3D interfaces.

reflected in the Table 5.1.

As we can observe in this Table, almost all users know about this topic due to this company constantly works in this field. It is important to highlight that this fact could have as many pros as cons. On the one hand, the users being aware of virtual environments can develop better their knowledge over the 3D virtual UIs and it could be easier for them to interact with these objects. Knowing about this topic, they could be more strict and critical allowing us to construct a more robust range of improvements for the future. On the other hand, the evaluation will not be so exhaustive for a person who has never interacted with a virtual UI limiting the opinions for this kind of person.

5.3 Demo Task

The task to evaluate the interaction with virtual UIs has been the Demo implemented and described in 4. The idea is that the users could advance along the different stages of the Demo without any external help.

As we will describe in 5.4, for the NASA-TLX model the user has to evaluate some dimension before start the Demo. After finishing the Demo the user has to complete two more forms to complete the mentioned NASA-TLX process and the QUE model.

5.4 Analysis and questionnaires

This section describes the subjective models used to evaluate the performance of our implemented UIs. As we have commented before, both models are based on fill out a form grading subjectively some questions. These two methods are described below.

5.4.1 NASA-TLX

NASA-TLX (explained in [26]) is a multidimensional assessment procedure that gives an overall workload score based on a weighted average of the scores in six sub-scales.

Dimension	Before Demo	After Demo
Mental demands	0-5	1-20
Physical demands	0-5	1-20
Temporal demands	0-5	1-20
Effort	0-5	1-20
Performance	0-5	1-20
Frustration level	0-5	1-20

Table 5.2: NASA TLX demands. An example form that a user receives to evaluate its Demo experience. Before testing the Demo the user has to evaluate the dimension in a one-five range. After testing it, the user evaluates the same dimensions but from one to twenty.

The real potential of this method, as mentioned in [28], the relevant factors which are relevant in the subjective experience of the workload are well defined being very clear to rate by the user.

The test protocol consists in two phases:

1. **Weighting phase**: the first one to determine the weights that the user gives to each dimension in a work task. The idea of this first evaluation is that the user gives his opinion about the importance of each dimension at the time of completing a certain work task before complete the new proposed task. Therefore, this first test is performed before the user completes the Demo. As we can observe in the Table 5.2, this assessment is scored between 0 and 5. As we will observe later these marks will be used to weigh the posterior evaluation in 2.

2. **Scoring phase**: the second one to evaluate each dimension according to their experience after completing a certain work task. The idea is to evaluate the previously scored dimensions again after testing the new task. This new assessment contributes to the opinion about the performance of the Demo once finished. As we can observe in the Table 5.2, this model is reported on a 20-point horizontal line scale. As we can observe down below these final scores along with the previous ones allow us to extract a final rate of the workload for the Demo task.

Therefore, the process to evaluate this method with the users has been the following. First, in the weighing phase, in which the user is asked about its experience in an interactive activity with virtual objects. In this way, the user has to evaluate from 0 to 5 the six dimensions of the Table 5.2. After that, the user test the Demo. Just after the user finished it, he is asked to evaluate his experience using the same previous dimensions but from 1 to 20. Finally, the total score is represented as a weighted sum of paired comparisons of the six dimensions.

Hence, NASA-TLX is a standardized test that provides information about the workload experienced by a user in several situations in which he has to meet the demands of a work task. Through this test, it has been possible to quantify objectively the interaction with virtual objects experience in terms of magnitude and importance as we will observe in the results section 5.5.

5.4.2 User Experience Questionnaire (UEQ)

UEQ is a quick evaluation to measure the user experience in an interactive task with any physical or virtual product as described in [27]. As we will see later, through this approach we can extract conclusions about our implemented UIs. This assessment is based also in six dimensions (as shown in the Table 5.3): attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. As exposed in the mentioned Table, these features are really descriptive about an interactive experience and include all the possible opinions that one user could have in our implemented Demo.

As described in [29], UEQ can be used to compare two interactive products. Given that the final stage of our Demo has two interfaces, one virtual and other physical, that perform the same activity, this evaluation will be really useful to evaluate in which aspects one UI is better than the other one.

Dimension	Description
Attractiveness	General impression about the UI
Perspicuity	Difficulty to get familiar with the UI
Efficiency	Is it necessary too much effort? Does the UI react fast?
Dependability	How are the UI control sensations? Security?
Stimulation	Motivating? Exciting?
Novelty	It is an innovative UI?

Table 5.3: UEQ dimensions. The UEQ method return the scores for six dimensions: attractiveness, perspicuity, efficiency, dependability, stimulation and novelty. These dimensions cover all the possible opinions that one user can have after testing the Demo.

These dimensions are computed by the questionnaire of 26 items in which the user has to evaluate several aspects of the tested UI. These items can be found in [27]. The table 5.4 shows a partial has been showed a partial example of the form that the user receives after completing the Demo. As we can observe, the user has to evaluate the different items from one to seven according to their sensations testing the Demo.

This form should be filled out spontaneously, without thinking too much the decision (this way we ensure that users report their real impression).

Annoying	1 2 3 4 5 6 7	Enjoyable
Dull	1 2 3 4 5 6 7	Creative
Boring	1 2 3 4 5 6 7	Exciting
...	1 2 3 4 5 6 7	...
Slow	1 2 3 4 5 6 7	Fast

Table 5.4: Some UEQ items. The example received by the user to evaluate his demo experience. The user has to evaluate from 1 to 20 the different items calculating eventually the results for the six described dimensions.

The process of this evaluation is simple:

1. **Explain the evaluation process**: before starting the Demo, the user is invited to focus on the difference between the physical and the virtual UIs. Given that the proposal of this evaluation is comparing both interactive sensations, we try to make the user aware of the importance of this task.

2. **The user test the Demo**: as we have commented before, in the virtual environment there are many indications about what to do in each stage. The user should be able to complete the Demo without asking how it works.

3. **Fill out the form**: the form is delivered to the user that should be completed quickly without analyzing too much its opinion in each item.

As we will observe in the result section 5.5, it will allow us to understand which aspects of our designed UIs are the best and which could be improved. As mentioned in [30], UEQ has been chosen as a standardized questionnaire showing a great acceptance in the scientific community for subjective analysis.

5.5 Results

This chapter presents the study of user opinions through the previously described evaluation methods. This analysis allows us to extract some interesting aspects of our implemented virtual UIs, as well as to compare the sensation among the virtual and physical UIs.

5.5.1 NASA-TLX results

Through the NASA-TLX evaluation, we have studied which dimensions assumed more importance for the users. For this assessment, each user has received a form as shown

	Mean weighted score
Mental demands	120
Physical demands	105
Temporal demands	176
Effort	132
Performance	180
Frustration level	123

Table 5.5: NASA-TLX results. On the one hand, the results show as the performance and temporal demands are the most voted. This means that, in spite of the task was easy, the users consider they have employed great efforts to complete the objective. On the other hand, nonetheless, the users say that they have not felt frustrated in the process.

in 5.2 that have filled out before and after completing the Demo.

The results obtained are shown in the Table 5.5. It has been calculated according to the NTP 544 protocol [1] and as shown in [31].

These values can give us which demands are highlighted in the user experience. As we can observe in the Table, the UI performance and temporal demands are the most voted dimensions. These dimension answers as shown in [31], to the question: what has been your satisfaction grade in the performance? what has been your pace of work? A high value means that the user had employed great efforts to complete the objective despite of it was an easy task. Also, they consider that the response of the UIs takes a long time. For these reasons, as we will comment in the future work chapter 6, these features will be real challenges to improve in the future.

5.5.2 UEQ results

Each user has been asked individually about each UI using the UEQ form. In this way, we have obtained four UEQ forms: three for the virtual UIs and one for the last physical UI. The results obtained have been analyzed separately addressing, finally, a comparison between the virtual and physical UIs.

5.5.2.1 Buttons panel UI

The first UEQ evaluation was focused on the Buttons panel UI described in 3.5.1. The results obtained can be observed in the Figure 5.1 in which are shown the six analyzed dimensions. In this representation, it is possible to observe two descriptive parameters for each dimension: the mean obtained for each one and the scores range

[1] https://www.insst.es

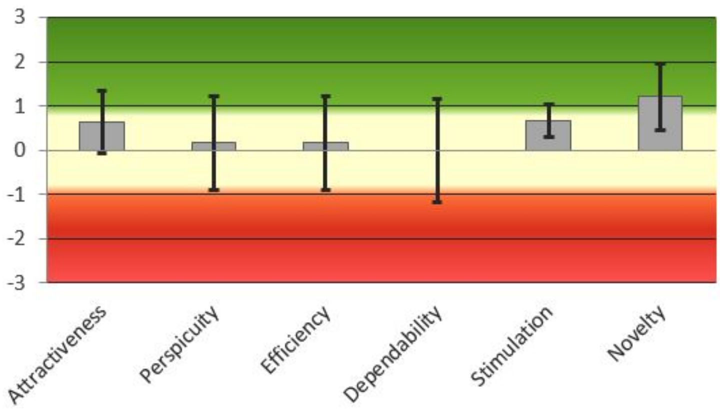

Figure 5.1: Buttons panel UI UEQ assessment. The opinion for this UI is distinguished by its novelty approach. The performance is acceptable but it should be improved to get a real amusing experience.

that the users have given to each one.

From this Figure, we could extract some conclusions:

- Pros: it has been an attractive and stimulating experience for the user. Above all, the users find it a novel concept of interaction.
- Cons: the users consider that the control sensations are not too enough accurate. Also, the opinions return that the UI reaction is not enough fast and that the needed effort should be lower.

5.5.2.2 Drop-down attached panel UI

The second evaluation was focused on the drop-down attached panel UI described in 3.5.2. The results are shown in the Figure 5.2 in which is collected the user opinions (as the previous evaluation). In general, the scores are better in all the dimensions being the more significant, the efficiency improvement. It is important to highlight how the users agree in their opinion being the score variations lower (black vertical lines).

From this Figure we could extract some conclusions:

- Pros: this UI is more attractive for the user than the previous one. Also, the interaction feels have improved significantly. It is considered also as a novel UI to experiment.
- Cons: the users keep thinking that is not easy to get familiar with the UI. The interaction feels are better but they should be improved.

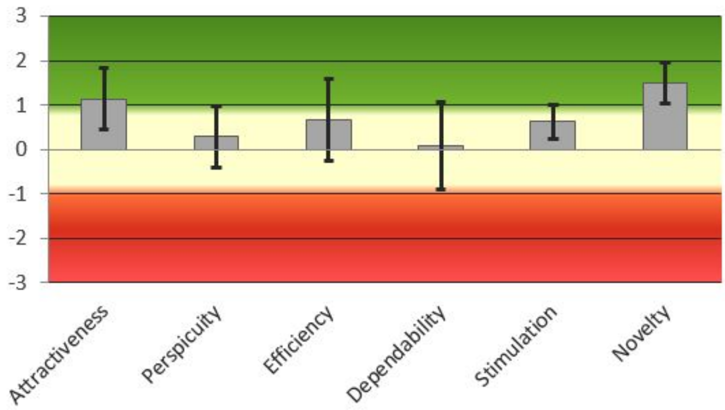

Figure 5.2: Drop-down attached panel UEQ assessment. The opinion for this UI is highlighted also by its novelty approach. The performance experience has improved significantly.

In the Demo, as we have commented in 4.1.4, this UI performs as a personal assistant in which the user can check how to continue if he does not know what to do. We could extract from the graphic that the user see this appreciation and consider it really attractive. However, they are critical to the correct performance of the interaction and consider that it should be improved.

5.5.2.3 Drop-down and movable panel UI

The third evaluation was focused on the drop-down and movable panel UI described in 3.5.3. We can observe the results in the Figure 5.3 in which is collected the scores given by the user for each dimension. At first sight, the opinions improve in some aspects, however, some of them, are worst getting lower scores than zero.

From this Figure we could extract some conclusions:

- Pros: the attractive, stimulation and novelty dimensions have high scores which means that the user positively the innovation approach of this UI.

- Cons: the perspicuity dimension has a really low score, which means that the efforts to get familiar with the UI are too high. The interact feelings should be improved to have a better experience.

In conclusion, we could say that this UI satisfies the attractive property being a really innovative experience for the user. The interact feelings could be also accepted, although the efforts to get familiar with the UI should be reduced.

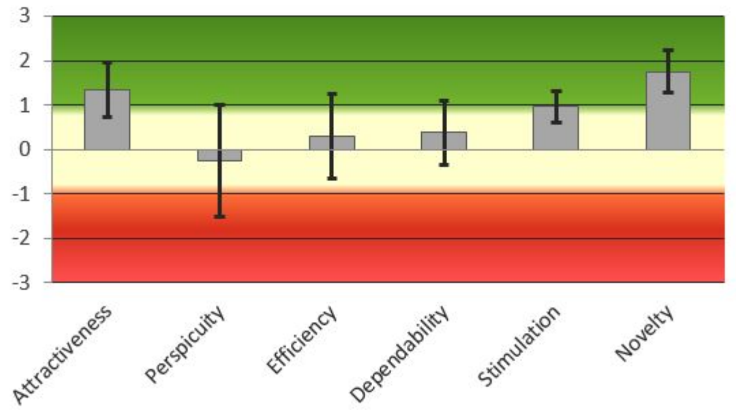

Figure 5.3: Drop-down and movable attached panel UEQ assessment. The opinion for this UI is highlighted once again by its attractiveness and novelty approach. In fact, these dimensions are really well evaluated. However, the performance experience has been reduced in several aspects.

5.5.2.4 Virtual UI vs. Physical UI

Finally, the last evaluation has been employed to compare the virtual UI with the physical UI through the UEQ form. As we have explained in 4.1.4, the final stage of the Demo has two interfaces, one virtual and other physical. These two interfaces make the same activity: manage a virtual TV. This, allows the user to test two kinds of interfaces in the same environment at the same time.

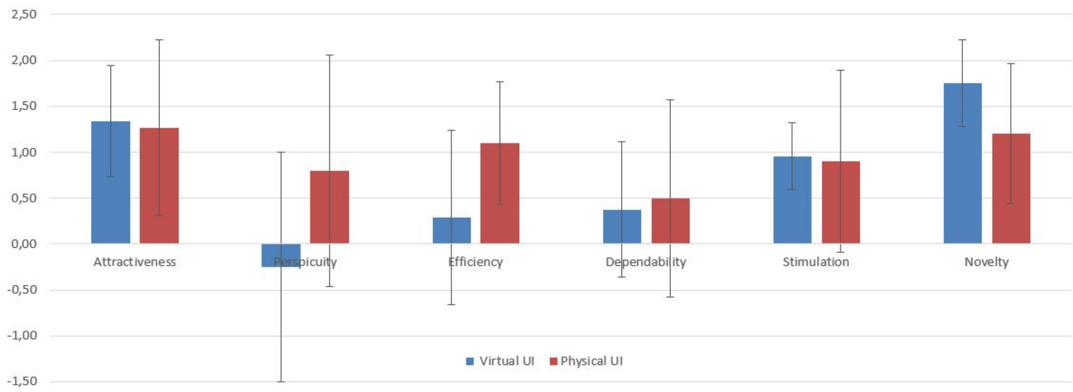

Figure 5.4: UEQ assessment for virtual and physical UIs. The physical UI highlight in its performance contributing to higher accuracy in the interaction skill. However, the novelty feelings and the attractive dimension is better scored for the virtual UI.

Using the UEQ evaluation we have compared the two opinions by the users about these interfaces. This, is shown in the Figure 5.4. In blue color is represented the scores about the virtual UI (already showed in 5.5.2.3) and in red the evaluation about the physical UI.

We can observe the opinion disparity in this graphic.

- **Virtual UI**: the users have determined that the interaction with virtual UIs is a real innovative sensation for them. The capacity to interact with virtual objects in a virtual environment observing their own hands increase the immersive sensation stunningly. However, the interaction task should be improved to get a fluent activity inside the virtual environment.

- **Physical UI**: the users have clear the better interact performance in the physical UI. Its experience touching the physic buttons presents in the tablet is more accurate for them. They have more security in the TV manage task taking a physical object in their hands. However, the feelings about this kind of interaction are not too exciting for them and it is less motivating than using a virtual UI.

Chapter 6

Conclusions and future work

6.1 Conclusions

Starting from the Distributed Reality idea proposed at the beginning, this document have contributed to study different approaches to create natural user interfaces for these kind of virtual environments. For that purpose, the proposed sub-objectives exposed in the introduction were achieved by the design and the implementation of different novel user interfaces. Subjective testing was conducted with several users to assess different usability parameters of these interfaces as other factors that impact the overall user experience.

An extensive state of the art analysis has been conducted involving most significant previous work in this field. This study has collected a great deal of information that has been the base for the implementation of our project.

The integration of the setup has been a true challenge. Starting from the setup composed of the HTC Vive HMD and the Zed Mini camera, it has been incorporated the Leap Motion device. The calibration in this combination has been the key to obtain a good performance. The design of the user interfaces has taken many resources in our work too. Leap Motion provides simple designs without any functionality. The logic performance of each UI has been a great challenge.

In order to create an attractive environment in which one user can test our implementation, we have designed a novel futuristic demo with the idea of creating an amusing experience for the user. We have tried to be original generating good-looking UIs to motivate the user in this experience.

We have used two subjective evaluation methods to evaluate the satisfaction grade of some users. These results have yielded important conclusions. Although the users consider these UIs as a novelty and attractive way of interaction, they think that the performance should be more fluent to get a really good experience.

Another interesting conclusion is that users still prefer physical interfaces and they find them more natural probably because they are used to interact mostly with physical environment. The lack of haptics feedback in the case of virtual interfaces makes them less appealing but they can be very useful as we demonstrated by the personal assistant.

6.2 Future work

As future work, we propose some improvements in order to resolve critical issues that we have found throughout the project. We could not fix these problems issues because most of them are related with combining several devices from different manufactures. For instance, the calibration is really difficult when integrating so many technologies in the same setup. We think that in the in the future we will have all this functionality integrated in the same device which will solve this issue or at least make it less cumbersome. Also, we propose some improvements to the future that could enhance the user experience using a virtual user interface.

- **Setup**: given that, as we have commented before, we have worked using three devices coming from three different companies, the integration has been a real challenge and we have had some problems. A few months ago, HTC has launched a new interesting HMD, the HTC Vive Pro[1]. This HMD contains two cameras simulating how the human vision works like the Zed Camera. This approach is really useful given that being part of the same setup it has no problems of integration. The accuracy between both hardware is perfect. Also, recent discoveries have worked in the hand tracking performance using these cameras through the disparity property (it is shown in the official HTC web page[2]).

- **Hands segmentation**: as we have explained, the hand's segmentation is implemented applying a chroma key. This task could be improved using another more complex segmentation techniques as it is exposed in [32]. Using these methods we could enhance the immersive experience.

- **UIs designed**: in this master's thesis, we have implemented three different UIs. In the future, it will be interesting to create new more complex UIs to cover a wider range of possibilities.

[1]https://www.vive.com/mx/product/vive-pro/
[2]https://developer.vive.com/resources/knowledgebase/vive-hand-tracking-sdk/

- **Evaluation**: it will be interesting to extend the evaluation process. There are other techniques of immersive evaluation like the IPQ questionnaire [33]. This assessment measures the immersion grade for one user within a virtual environment reporting other important opinions that we could consider to improve the experience.

Bibliography

[1] Z. Zhang, B. Cao, J. Guo, D. Weng, Y. Liu, and Y. Wang, "Inverse virtual reality: Intelligence-driven mutually mirrored world," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 735–736, March 2018.

[2] E. G.-S. A.Villegas, P.Perez, "Dr," in *Towards a Distributed Reality: a Multi Video approach to Xr*, Proc. ACM 11th International Workshops on Immersive Mixed and Virtual Environment Systems, 2019.

[3] C. Carlsson and O. Hagsand, "Divea platform for multi-user virtual environments," *Computers Graphics*, vol. 17, no. 6, pp. 663 – 669, 1993.

[4] J. Pejoska, M. Bauters, J. Purma, and T. Leinonen, "Social augmented reality: Enhancing context-dependent communication and informal learning at work," *British Journal of Educational Technology*, vol. 47, no. 3, pp. 474–483, 2016.

[5] S. Su, V. Perry, H. Roy, K. Gamble, and S. Kase, "3d user interface for a multi-user augmented reality mission planning application," in *Advances in Human Factors and Simulation* (D. N. Cassenti, ed.), (Cham), pp. 120–131, Springer International Publishing, 2020.

[6] A. Nowak, M. Woźniak, M. Pieprzowski, and A. Romanowski, "Advancements in medical practice using mixed reality technology," in *Innovations in Bio-Inspired Computing and Applications* (A. Abraham, N. Gandhi, and M. Pant, eds.), (Cham), pp. 431–439, Springer International Publishing, 2019.

[7] R. K. J. R. I. B. F. P. P. Perez, E. Gonzalez-Sosa and A. Villegas, "Dr," in *Immersive Gastronomic Experience with Distributed Reality*, In Proc. IEEE 5th Workshop on Everyday Virtual Reality (WEVR), 2019.

[8] C.-I. Chao, C.-M. Wang, H.-C. Kuo, L.-C. Tseng, S.-K. Lin, Y.-J. Tsai, C.-C. Lin, and D.-F. Chang, "Live room merger: A real-time augmented reality system for merging two room scenes," in *Proceedings of the Virtual Reality International*

Conference - Laval Virtual 2017, VRIC '17, (New York, NY, USA), pp. 8:1–8:4, ACM, 2017.

[9] H. H. Baker, N. Bhatti, D. Tanguay, I. Sobel, D. Gelb, M. E. Goss, W. B. Culbertson, and T. Malzbender, "Understanding performance in coliseum, an immersive videoconferencing system," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 1, pp. 190–210, May 2005.

[10] C. Anthes, R. J. GarcÃa-HernÃ¡ndez, M. Wiedemann, and D. KranzlmÃŒller, "State of the art of virtual reality technology," in *2016 IEEE Aerospace Conference*, pp. 1–19, March 2016.

[11] A. Borrego, J. Latorre, M. Alcaniz, and R. Llorens, "Comparison of oculus rift and htc vive: Feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation," *Games for Health Journal*, vol. 7, no. 3, 2018.

[12] M. L. Diederick C. Niehorster, Li Li, "The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research," *GH*, 2017.

[13] G. Bruder, F. Steinicke, K. Rothaus, and K. Hinrichs, "Hm," in *Enhancing Presence in Head-Mounted Display Environments by Visual Body Feedback Using Head-Mounted Cameras*, pp. 43–50, 01 2009.

[14] Z. Ma, H. Huang, X. Yang, and B. Li, "Pose estimation for unknown motion target based on improved icp algorithm and adaptive ekf," in *2018 IEEE International Conference on Cyborg and Bionic Systems (CBS)*, pp. 106–111, Oct 2018.

[15] T. Vodopivec, V. Lepetit, and P. Peer, "Fine hand segmentation using convolutional neural networks," *CoRR*, vol. abs/1608.07454, 2016.

[16] M. Oberweger, P. Wohlhart, and V. Lepetit, "Hands deep in deep learning for hand pose estimation," *CoRR*, vol. abs/1502.06807, 2015.

[17] M. Khademi, H. Mousavi Hondori, A. McKenzie, L. Dodakian, C. V. Lopes, and S. C. Cramer, "Free-hand interaction with leap motion controller for stroke rehabilitation," in *Proceedings of the Extended Abstracts of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI EA '14, (New York, NY, USA), pp. 1663–1668, ACM, 2014.

[18] B. D. Weichert Frank, “Lm,” in *Analysis of the accuracy and robustness of the leap motion controller*, 2013.

[19] S. Bodker, “U,” in *Through the Interface - a Human Activity Approach to User Interface Design*, 1991.

[20] E. K. Dog Bowman, Joseph J. LaViola Jr., “U3,” in *3D User Interfaces: Theory and Practice*, 2004.

[21] P. NorthStar, “Lm,” in *"The world’s most powerful hand tracking is here" Leap Motion’s Unity SDK 4.4.0*, 2018.

[22] D. King, M. Horton, and D. R. Lamichhane, “Evaluating the usability of the leap motion controller,” in *Proceedings of the 32Nd International BCS Human Computer Interaction Conference*, HCI ’18, (Swindon, UK), pp. 183:1–183:5, BCS Learning & Development Ltd., 2018.

[23] M. Suvajdzic, A. Bihorac, P. Rashidi, T. Ong, and J. Applebaum, “Virtual reality and human consciousness: The use of immersive environments in delirium therapy,” *Technoetic Arts*, vol. 16, pp. 75–83, 03 2018.

[24] J. Jung, H. Lee, J. Choi, A. Nanda, U. Gruenefeld, T. Stratmann, and W. Heuten, “Ensuring safety in augmented reality from trade-off between immersion and situation awareness,” in *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 70–79, Oct 2018.

[25] D. Anton, G. Kurillo, and R. Bajcsy, “User experience and interaction performance in 2d/3d telecollaboration,” *Future Generation Computer Systems*, vol. 82, pp. 77 – 88, 2018.

[26] E. Diaz Ramiro, S. Rubio Valdehita, J. Marta Garcia, and L. Lucia Moreno, “Estudio Psicom Indice de Carga Mental NASA-TLX con una Muestra de Trabajadores Espanoles,” *Revista de Psicologia del Trabajo y de las Organizaciones*, vol. 26, pp. 191 – 199, 12 2010.

[27] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Design and evaluation of a short version of the user experience questionnaire (ueq-s),” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, p. 103, 01 2017.

[28] C. Nikulin, G. Lopez, E. Piñonez, L. Gonzalez, and P. Zapata, “Nasa-tlx for predictability and measurability of instructional design models: case study in design methods,” *Educational Technology Research and Development*, vol. 67, pp. 467–493, Apr 2019.

[29] M. Rauschenberger, M. Schrepp, M. Cota, S. Olschner, and J. Thomaschewski, "Efficient measurement of the user experience of interactive products. how to use the user experience questionnaire (ueq). example: Spanish language version," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 2, pp. 39–45, 03 2013.

[30] M. Schrepp, A. Hinderks, and J. Thomaschewski, "Design and evaluation of a short version of the user experience questionnaire (ueq-s)," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 4, p. 103, 01 2017.

[31] E. Galy, J. Paxion, and C. Berthelon, "Measuring mental workload with the nasa-tlx needs to examine each dimension rather than relying on the global score: an example with driving," *Ergonomics*, vol. 61, no. 4, pp. 517–527, 2018. PMID: 28817353.

[32] B. Kang, K. Tan, N. Jiang, H. Tai, D. Tretter, and T. Nguyen, "Hand segmentation for hand-object interaction from depth map," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 259–263, Nov 2017.

[33] H. Regenbrecht and T. Schubert, "Real and illusory interactions enhance presence in virtual environments," *Presence: Teleoper. Virtual Environ.*, vol. 11, pp. 425–434, Aug. 2002.